

RESEARCH

Open Access

# An inverse metabolic engineering approach for the design of an improved host platform for over-expression of recombinant proteins in *Escherichia coli*

Chaitali Ghosh, Rashmi Gupta and Krishna Jyoti Mukherjee\*

## Abstract

**Background:** A useful goal for metabolic engineering would be to generate non-growing but metabolically active quiescent cells which would divert the metabolic fluxes towards product formation rather than growth. However, for products like recombinant proteins, which are intricately coupled to the growth process it is difficult to identify the genes that need to be knocked-out/knocked-in to get this desired phenotype. To circumvent this we adopted an inverse metabolic engineering strategy which would screen for the desired phenotype and thus help in the identification of genetic targets which need to be modified to get overproducers of recombinant protein. Such quiescent cells would obviate the need for high cell density cultures and increase the operational life span of bioprocesses.

**Results:** A novel strategy for generating a library, consisting of randomly down regulated metabolic pathways in *E. coli* was designed by cloning small genomic DNA fragments in expression vectors. Some of these DNA fragments got inserted in the reverse orientation thereby generating anti-sense RNA upon induction. These anti-sense fragments would hybridize to the sense mRNA of specific genes leading to gene 'silencing'. This library was first screened for slow growth phenotype and subsequently for enhanced over-expression ability. Using Green Fluorescent Protein (GFP) as a reporter protein on second plasmid, we were able to identify metabolic blocks which led to significant increase in expression levels. Thus down-regulating the *ribB* gene (3, 4 dihydroxy-2-butanone-4-phosphate synthase) led to a 7 fold increase in specific product yields while down regulating the gene *kdpD* (histidine kinase) led to 3.2 fold increase in specific yields.

**Conclusion:** We have designed a high throughput screening approach which is a useful tool in the repertoire of reverse metabolic engineering strategies for the generation of improved hosts for recombinant protein expression.

**Keywords:** Recombinant protein, Inverse metabolic engineering, *Escherichia coli*, Improved host platform

## Background

Improvement of metabolic phenotype through directed genetic modifications is the main goal of metabolic engineering [1-4]. The classical approach of metabolic engineering requires a detailed knowledge of enzyme kinetics, the system network, and intermediate pools involved, and on this basis, a genetic manipulation is proposed for some presumed benefits. Recent progress in molecular genetics methods makes it possible to

knockout or over-express targeted genes in most microorganisms [5-9]. A key question in metabolic engineering is how to identify gene targets that have direct or indirect impact on a particular phenotype of interest [10].

For the over-production of metabolites the issue is comparatively straight forward since only the regulatory blocks and bottlenecks in the specific pathway involved in product synthesis need to be removed. Secondly, the supply of precursor metabolites to the pathway need to be enhanced thereby improving the metabolic flux in the pathways. However, for products like recombinant

\* Correspondence: [kjmukherjee@mail.jnu.ac.in](mailto:kjmukherjee@mail.jnu.ac.in)  
School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

proteins there is a vast range of precursors along with the transcriptional/translational mechanism all of which are coupled to the growth process. Many metabolic engineering approaches have therefore attempted to enhance both growth and recombinant protein production in *E. coli* [11,12]. One approach has been to supplement the genes which get down-regulated due to the stress associated with recombinant protein expression [13]. Similarly key factors which has impact on growth and expression yield, like glycolysis and Tri Carboxylic Acid (TCA) cycle enzymes, ATPase's, RNA polymerases etc. have been co-expressed [14-16].

A novel alternative strategy to design overproducers of the recombinant product would be to make these precursors available specifically for product rather than biomass formation. We therefore decided to design a strategy to screen for metabolically active but non-growing quiescent cells which have been earlier shown to be more effective for recombinant protein production [17].

The concept of inverse metabolic engineering involves first to identify the desired phenotype, then to determine environmental or genetic conditions that confer this phenotype, and finally to alter the phenotype of the selected host by genetic manipulation [18,19]. In this study, we adopted this approach where an *E. coli* genomic library was screened to obtain quiescent cells which might be better producers of recombinant proteins. For this, we randomly and partially down-regulated the biosynthetic pathways in the cell, in order to identify those blockages which help in diverting the metabolic flux away from growth. Cells with slow growth phenotype or growth stoppage were then screened for enhanced protein expression capability. An additional advantage of non-growing, quiescent cells is that they can increase the operational life span of bio-processes and improve process economics by decoupling product formation from cell growth [20].

In this study an *E. coli* genomic library was prepared in a pRSET A vector having a strong promoter as well as in pBAD33 having a comparatively weaker promoter because even a small down regulation of some genes can have significant effects. Screening of the genomic library by different approaches (slow growth or enhanced GFP fluorescence) led to the identification of clones with different down-regulated genes which blocked growth and simultaneously improved recombinant protein expression.

## Results and discussion

### Identification of metabolic pathways which lead to 'no growth' or slow-growth phenotype through genomic library screening

We constructed a library consisting of a set of clones with metabolic pathways which were 'knocked-down' rather than 'knocked-out,' allowing us to include the role of essential genes. In our study the strategy adopted was

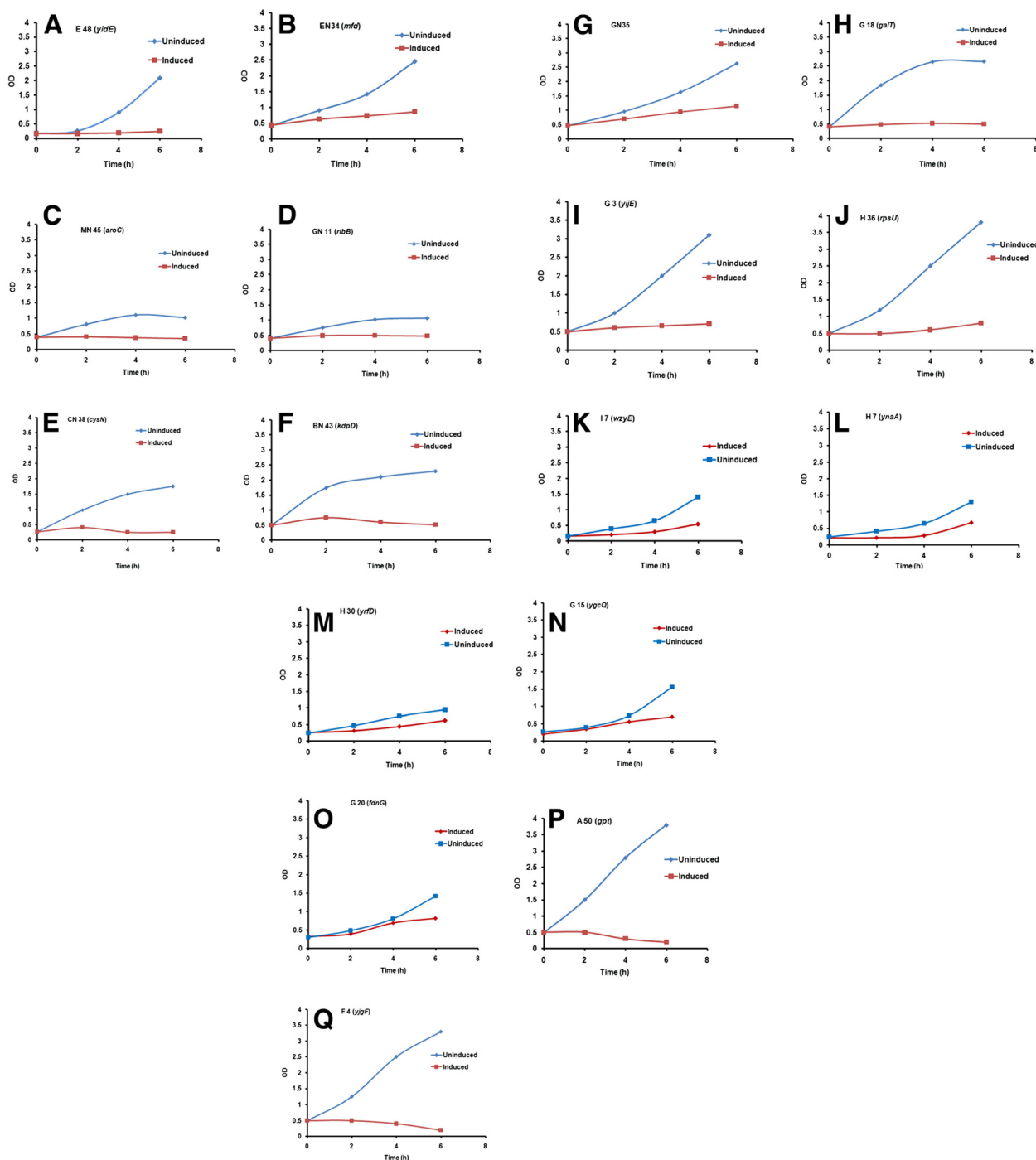
to clone small fragments of genomic DNA (~ 200–800 bases), the idea being that some of these DNA fragments would get inserted in an opposite orientation (compared to the coding strand) in the expression vector. On induction, the RNA produced would be complimentary to the mRNA (of a functional protein) which would then hybridize leading to (partial) silencing of the gene. In eukaryotic systems an ideal antisense could be as short as a 22–25 bp fragment (as then a dicer molecule can act, leading to RNAi). But in prokaryotic systems, RNAi is absent and not many studies have been done on the antisense mechanism. So by choosing a 200–800 bp fragment we attempted to ensure that at least partial gene silencing would take place.

The concept of inverse metabolic engineering was used to screen this library for the desired 'no-growth' phenotype and then to identify the genetic factors that confer this phenotype. We were interested in a 'no-growth' or 'slow-growth' phenotype where the metabolic activity was undiminished; the rationale being such clones would divert the metabolic flux away from biomass formation and towards product synthesis.

### Construction of the antisense library in pRSET A vector

Genomic DNA fragments of sizes ranging from 200–800 bp (as we required only a part of the gene which would act in an antisense fashion to block transcription) were isolated after an optimized partial digestion of the *E. coli* genomic DNA and subsequently ligated into a high copy number vector (pRSET A) with a strong T7 promoter. The constructed shotgun library in pRSET A was transformed into *E. coli* BL21 pLysS strain.

A two step screening strategy was used. In the first step, more than 8000 transformants were screened. Preliminary screening was done on plates. Each transformant was replica plated and only those clones were selected which upon Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) induction grew very slowly on plates. After first screening, 728 clones were isolated and a second round of screening was done on LB agar plates. Finally 70 clones were picked and the metabolic activity of these slow growing transformants were checked in shake flask cultures containing M9 media (supplemented with glucose), by looking at rates of decline of glucose in the medium, post IPTG induction. Our aim was to identify clones with negligible growth rate but unimpaired metabolic activity. 17 transformants were finally selected with this desirable phenotype where it was expected that the metabolic flux would possibly get diverted towards recombinant protein expression instead of growth. Shake flask studies of all these 17 clones were done in different media (LB, TB & M9) where a similar retardation in growth was observed. As is clear from Figure 1 significant growth retardation was observed post induction compared to the uninduced cultures.



**Figure 1** Growth profiles of the clones in LB media under T7 promoter of pRSET A, in the host BL21 pLysS. (A) E48 (*yidE*); (B) EN34 (*mfd*); (C) MN45 (*aroC*); (D) GN11 (*ribB*); (E) CN38 (*cysA*); and (F) BN43 (*kdpD*) (G) GN35; (H) G18 (*galT*); (I) G3 (*yjiE*); (J) H36 (*rpsU*); (K) I7 (*wzyE*); and (L) H7 (*ynaA*) (M) H30 (*yrfD*); (N) G15 (*ygcQ*); (O) G20 (*fdnG*); (P) A50 (*gpt*) and; and (Q) F4 (*yjfF*) under (◻) uninduced; and (◼) IPTG induced condition respectively.

These clones were sequenced and used to do a BLAST search against the *E. coli* genome database. The lists of identified genes are given in Table 1. The details of the genes were collected from the PEC (Profiling of *E. coli* Chromosome) and Ecocyc database.

These 17 clones were co-transformed with a second plasmid pBAD33-GFP, where the GFP gene was cloned under the arabinose promoter and had a compatible p15 'ori' with pUC 'ori' of pRSET A. This GFP reporter gene was used to check whether the antisense expression (while

**Table 1 List of transcripts whose blockage leads to growth stoppage**

Transcript name	Gene name	Size (bp)	Known or putative function of the gene product	Orientation	Essential/ Non essential
GN35 (T7)	-	221	Ketoacid-binding protein	Reverse	Non-essential
CN38 (T7)	<i>cysN</i>	567	Adenosine 5' phosphosulphate kinase sulfate adenylyltransferase subunit 1	Same direction	Non-essential
BN43 (T7)	<i>kdpD</i>	568	Sensory histidine kinase	Reverse	Non-essential
GN11 (T7)	<i>ribB</i>	403	3,4 dihydroxy-2-butanone-4-phosphate synthase	Reverse	Non-essential
EN34 (T7)	<i>mfd</i>	587	Transcription repair coupling factor	Same direction	Non-essential
MN45 (T7)	<i>aroC</i>	190	Chorismate synthase	Reverse	Non-essential
G3 (T7)	<i>yjiE</i>	379	Predicted permease	Reverse	Non-essential
A50 (T7)	<i>gpt</i>	141	Xanthine-guanine phosphoribosyl transferase	Reverse	Non-essential
F4 (T7)	<i>yjgF</i>	122	Ketoacid binding protein	Reverse	Non-essential
H36 (T7)	<i>rpsU</i>	422	Acid resistant protein	Same direction	Non-essential
E48 (T7)	<i>yidE</i>	350	Predicted transporter	Reverse	Non-essential
G15 (T7)	<i>ycgQ</i>		Predicted electron transfer flavoprotein, NAD/FAD binding	Same direction	Non-essential
G18 (T7)	<i>galT</i>	375	Galactose -1 phosphate uridylyltransferase	Reverse	Non-essential
G20 (T7)	<i>fdnG</i>	720	Formate dehydrogenase- N, alpha subunit, nitrate inducible	Reverse	Non-essential
H7 (T7)	<i>ynaA</i>	876	Conserved hypothetical protein: Rac prophage	Reverse	Non-essential
H30 (T7)	<i>yrfD</i>	769		Reverse	Non-essential
I7 (T7)	<i>wzyE</i>	770	Predicted Wzy protein involved in ECA polysaccharide	Same direction	Non-essential
L27 (ara)	-	639	Conserved hypothetical protein	Reverse	Non-essential
A17 (ara)	<i>bioF</i>	414	Predicted methyltransferase, enzyme of biotin synthesis	Reverse	Non-essential
A29 (ara)	<i>kdpF</i>	289	2-octaprenylphenol hydroxylase enzyme	Reverse	Non-essential
F50 (ara)	-	313	iro p300	Reverse	Non-essential

blocking growth) also helped in the overexpression of recombinant protein. All these experiments were done in LB media. For all the 17 clones one flask was induced only with 0.02% of arabinose (control), while the second test flask was co-induced with both IPTG and arabinose, to check the effect of the antisense transcript on GFP expression. A third flask for all the co-transformants was kept uninduced to observe the effect if any, of leaky expression (2<sup>nd</sup> control). The library contained large number of random inserts in pRSET A (which did not show any effect) which essentially served as a negative control.

Figure 2 (A,B,C) shows the effect of co-expression of clone GN11 having the anti-sense of *ribB* gene (*rib-3, 4* dihydroxy-2-butanone-4-phosphate synthase). In Figure 2B the uninduced culture had a basal level GFP expression of 5 to 7 AU while GFP expression in the *ara* induced control culture increased linearly from 6 to 39 AU in 8 hours post induction. In the test system the fluorescence of the co-induced culture overtook the control flask after 2 hours to give a final fluorescence of 53 AU demonstrating the positive effect of antisense induction on GFP expression. Since the final biomass concentrations obtained in the test culture was significantly lower, the specific product yield was much higher

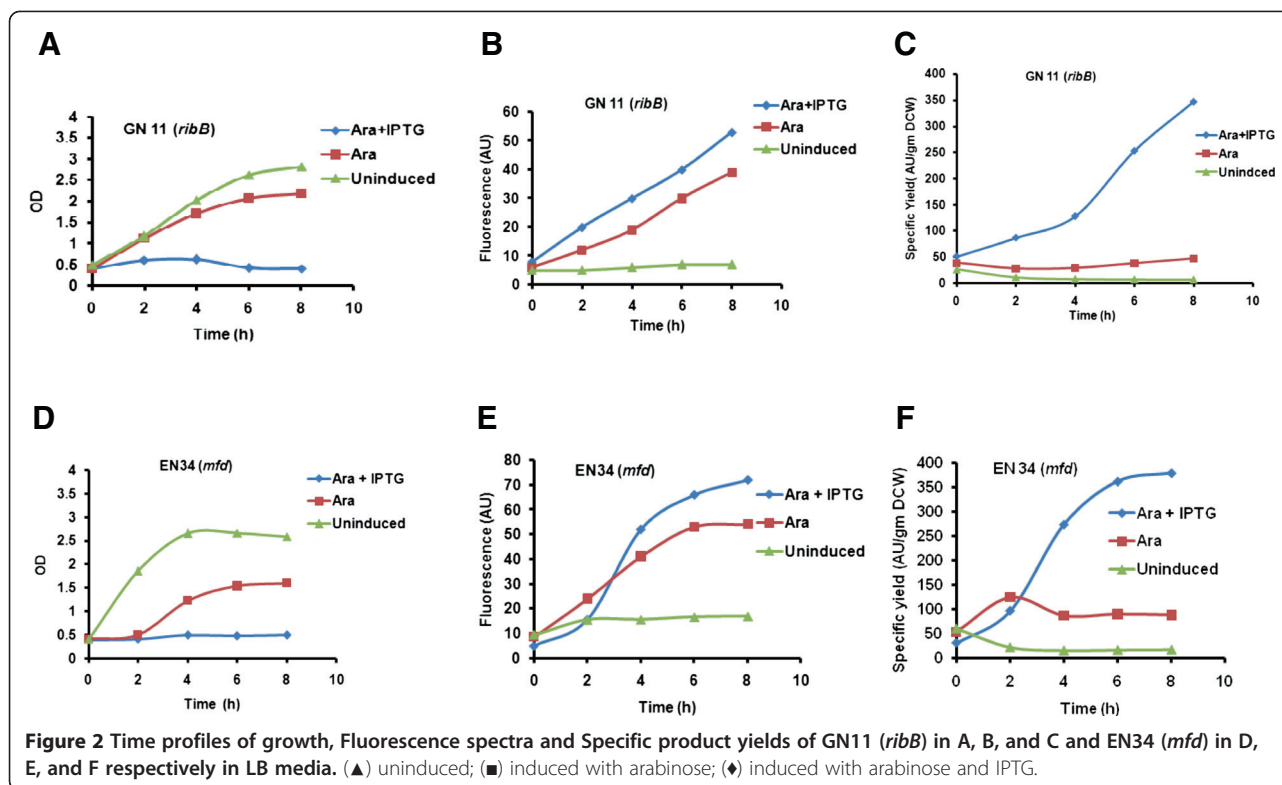
at 347 AU/g DCW, a 7 fold increase over the control culture (47 AU/g DCW) (Figure 2C).

Figure 2 (D,E,F) shows the co-expression of the clone EN34 coding for mutation frequency decline (*mfd*) protein. In Figure 2E the uninduced flask showed a basal level expression of 16.9 AU throughout the cultivation. The control flask showed a linear increase from 10 AU to 54 AU in 8 hours. However, while the expression was initially low in the co-induced culture, it overtook the control flask after 3 hours and gave a final GFP fluorescence of 72 AU. The specific product yield of arabinose induced GFP was found to be 88 AU/gm DCW in the 8<sup>th</sup> hour post induction in the control culture whereas in the arabinose and IPTG induced culture it was calculated to be 378 AU/gm DCW which was 4 fold higher (Figure 2F).

Similar studies were conducted with the rest of the 15 clones. However, not all the clones gave such a dramatic increase in the specific product yield though we did observe a significant increase in specific product yield in some cases which are listed in Table 2.

#### Construction of the antisense library in pBAD33

In the previous screening strategy we over-expressed the transcripts under the strong T7 promoter, which has



certain disadvantages. Firstly, IPTG itself is partially toxic to cells and also the expression of T7 RNA polymerase leads to growth retardation [21,22]. This would mask the true growth retardation due to the anti-sense effect of the transcripts. Secondly, since the anti-sense works at the RNA

**Table 2** Specific product yield (AU/gm DCW) of various clones under three different conditions, uninduced, induced with arabinose and induced with arabinose and IPTG both (6 hrs post induction)

a) Clones in pRSET A			
Clone name	Specific product yield (Ara + IPTG)	Specific product yield (Ara)	Specific product yield (Uninduced)
EN34 ( <i>mfd</i> )	362	91	14
G18 ( <i>galT</i> )	76	227	18
GN35	89	244	9
GN11 ( <i>ribB</i> )	347	47	8
MN45 ( <i>aroC</i> )	67	199	25
BN43 ( <i>kdpD</i> )	256	67	10
b) Clones in pBAD 33			
A17 ( <i>bioF</i> )	125	181	9
A29 ( <i>kdpF</i> )	681	79	5
F50	225	207	10
L27	175	175	13

level we should use only a transcriptional vector and not an expression vector to rule out the possibility that the toxicity of a randomly expressed polypeptide is responsible for growth retardation. Thirdly, hyper expression of the anti-sense RNA seems to be a clumsy approach to block and/or down-regulate gene expression (even though this was primarily a screening strategy for identifying potential targets for manipulation). We therefore chose a much better regulated transcriptional vector for the construction of a second library which would identify targets where even a partial down-regulation would lead to diversion of metabolic flux towards product synthesis.

The genomic library was prepared under the *ara* promoter in pBAD33 (a transcriptional vector), which is a tightly regulated promoter. Moreover as we were not constrained to use the BL21 (DE3) strain we chose *E. coli* DH5 $\alpha$  since it has superior transformation efficiency and is much more stable strain for retaining plasmids. A dual plasmid system was used to directly screen for the desired phenotype. Competent DH5 $\alpha$  cells pre-transformed with a plasmid pNER31 (having GFP under the *lac* promoter) was used for co-transformation with the genomic library. This allowed us to directly screen for the over-expressing phenotype. Upon co-induction of the transcripts, we screened for colonies exhibiting intense green color of GFP (under UV) compared to the control colonies in the plate. More than 30,000

transformants of the library was screened. Finally 4 clones (Table 1), exhibiting higher GFP expression was selected for further studies. Shake flask studies on these clones were done in different media LB, TB, M9 along with appropriate controls.

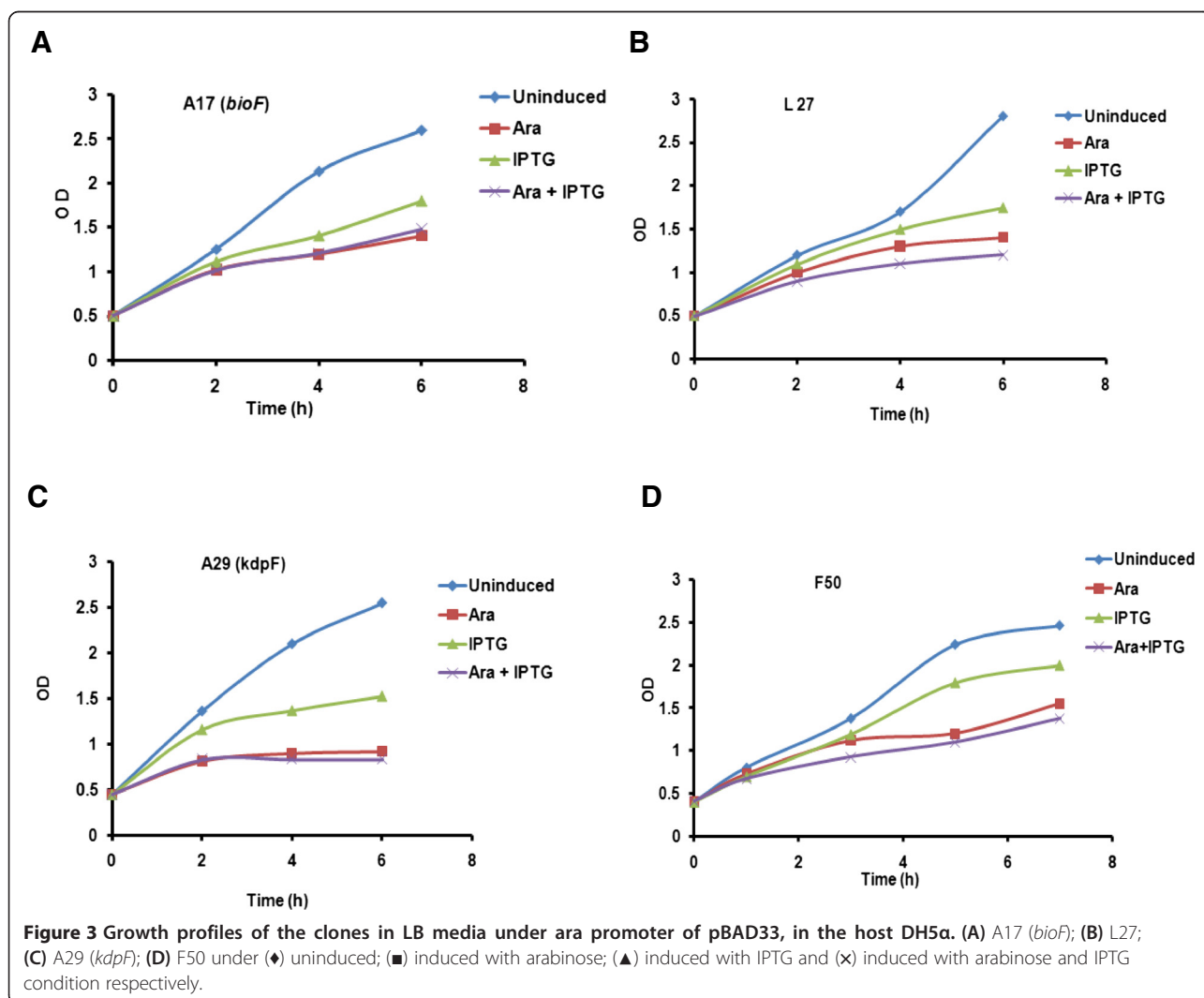
We first needed to factor out the effect of GFP expression on growth retardation by comparing the growth of the uninduced culture with the IPTG induced culture. A very slight decline in growth was observed post induction in LB media which is typical of GFP induced cultures. However when the second, anti-sense gene carrying plasmid, was induced by arabinose a significant drop in growth was observed. The growth profile of the co-induced culture was compared with the control culture where only GFP was induced. We observed that the growth declined sharply in all the co-induced cultures (Figure 3).

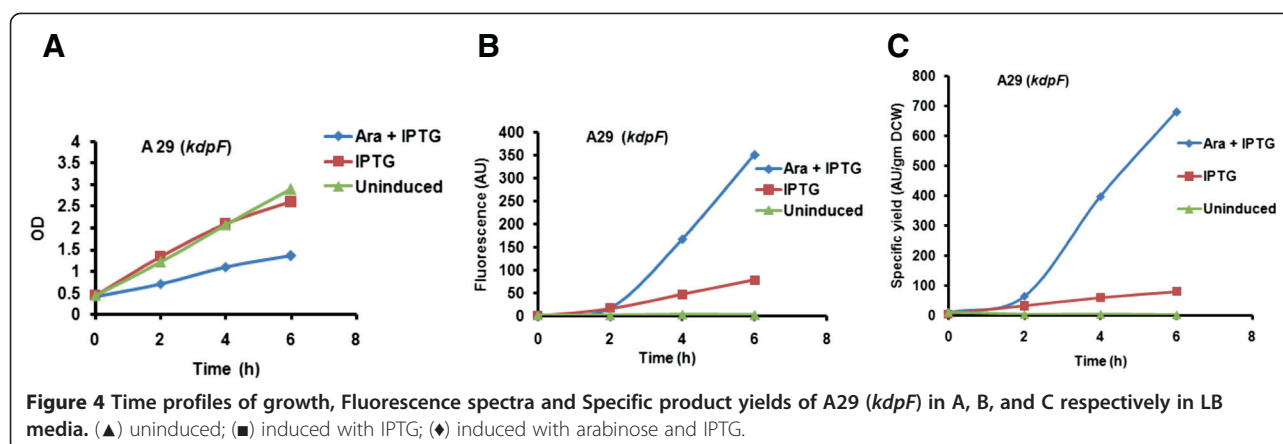
Comparing the product profile of *kdpF* (2-octaprenyl-phenol hydroxylase enzyme) antisense expressing clone A29, we observed that, the specific product yield of the

co-induced culture was 8.4 fold higher than the control culture which was induced with IPTG only (Figure 4). Co-expression studies were also done with rest of the 3 clones (data not shown).

One of the open questions in the quest for improved host platforms is how targets are to be identified which confer the desired phenotype, given the lack of a comprehensive kinetic model for *E. coli* and the absence of detailed information for the regulatory mechanism, which operate inside the cells. There are therefore several approaches for the modifications of the *E. coli* host to achieve the desired goal [23-25]. In our case the design strategy was essentially a high throughput screen which helped us select better producers. It is interesting to note that all the selected genes were non-essential and also not directly linked to either growth or product formation.

For example, the selected clone GN11 contained an insert of 403 bp of the *ribB* gene (3, 4 dihydroxy-2-butanone-4-phosphate synthase) in the reverse direction.





*ribB* serves as the biosynthetic precursor for the xylene ring of riboflavin [26,27]. It is reported that *ribB* mutant *E. coli* leads to slow growth but remains metabolically active which corroborates our studies [28]. However the discovery that blocking the gene *ribB* can help divert the metabolic flux towards recombinant protein production was both non obvious and counter intuitive. Similarly, clone BN43 coded for a protein KdpD (sensory histidine kinase) which is expressed during osmotic shock and is otherwise nonessential. Clearly in this case there is no direct link to the gene function and the observed channelization of metabolic flux towards recombinant protein production. In the case of clone MN45 which contained part of the antisense *aroC* (chorismate synthase gene, which catalyzes the formation of chorismate in aromatic amino acid biosynthesis) its blockage should ideally lead to a decrease in the biosynthetic pathways involved in growth. However, an exact estimate of the growth inhibition is difficult given that complex nitrogen sources are present in the medium which can supply the growing cell with the desired amino acids. The fact that blocking chorismate synthase led to a complete growth stoppage was interesting and points to the criticality of even single metabolic reactions in growth. All these results underscore the fact that there is no direct link at the pathway level between gene function and observed phenotype which could possibly be the result of a complex regulatory response within the cell.

## Conclusion

In a novel approach towards the design of an ideal host platform for the over-expression of recombinant proteins, two different high throughput screening strategies were designed to identify genes whose down-regulation would lead to a slow growth phenotype and also possibly divert the metabolic flux towards over-expression of recombinant proteins. The leads obtained did not rely on a *priori* knowledge of the regulated and interconnected nature of the *E. coli*'s metabolic network. At a practical

level we were able to obtain very useful targets for gene knock-out/knock-down which would enable the design of better hosts for protein expression. It would be interesting to study the combined and synergistic effect of these identified blocks. At a more fundamental level a detailed analysis of the phenotypic affects of these metabolic blocks can lead to a better understanding of the regulatory mechanisms in the *E. coli* network.

## Materials and methods

### Growth media and antibiotics

Media and bulk chemicals were purchased from local manufacturers, Himedia, Qualigens, Difco and MERCK. Media used were Luria Broth [LB] (bactotryptone 10 g/l, yeast extract 24 g/l and NaCl 15 g/l), Terrific Broth [TB] (YE 24 g/l, tryptone 12 g/l, glucose 0.5% in which 2.31 g/l  $\text{KH}_2\text{PO}_4$  and 12.54 g/l  $\text{K}_2\text{HPO}_4$  were added after autoclaving separately), and minimal media [M9] (NaCl 0.5 g/l,  $\text{NH}_4\text{Cl}$  1 g/l,  $\text{K}_2\text{HPO}_4$  3 g/l, 1 M  $\text{CaCl}_2$  0.1 ml/l, 1 M  $\text{MgSO}_4$  2 ml/l and 20% glucose 10 ml/l). Antibiotics used were ampicillin 100  $\mu\text{g/ml}$  (1X), kanamycin 50  $\mu\text{g/ml}$  (1X) and chloramphenicol 30  $\mu\text{g/ml}$  (1X).

### Bacterial strains and plasmids

The T7 based expression vector pRSET A (ampicillin, Invitrogen) which had pUC 'ori', was used in these studies. pBAD33 (chloramphenicol, Beckwith) [29] having p15 'ori' with *ara* promoter and compatible with pRSET A, was used for expression of GFP gene (used as a reporter protein) and the construction of second library. The plasmids were maintained and expressed in *E. coli* DH5 $\alpha$  cells. *E. coli* BL21 (DE3) from stratagene (USA) and *E. coli* BL21 (DE3) pLysS were used for expression studies. Strain JM109 (Promega, USA) was used for the genomic library preparation. pNER31 (ampicillin) having GFP gene under '*lac*' promoter and induced by IPTG, was gifted by Summers, Cambridge, UK and pET14b-GFP (ampicillin) where GFP gene was cloned under T7 promoter, was gifted by J. Singh, ICGEB, India.

### Construction of genomic library in vectors pRSET A and pBAD33

Genomic DNA was isolated from JM109 by the Hexadecyltrimethylammonium bromide (CTAB) method. This was partially digested with either *Sau*III A or *Nla*III both four base cutters. The expression vectors pRSET A and pBAD33 were digested with *Bgl*II and *Sph*I which generate compatible cohesive ends with *Sau*III A and *Nla*III respectively. These linearised vector fragments were Calf Intestinal Alkaline Phosphatase (CIAP) treated to prevent self ligation and ligated with the digested and gel eluted genomic fragments of size ranging from 200 to 800 bp. The ligation mixture was used to transform electrocompetent BL21(DE3) pLysS and DH5 $\alpha$  cells respectively and plates were incubated at 37°C overnight. Approximately 8000 and 30,000 colonies were obtained in each library respectively. The quality of the library that was constructed was checked by picking a few colonies and checking for the insert size by colony PCR.

### Screening of the library by two different approaches

A preliminary screening of the pRSET A vector library was done on LB agar plates. Replica plating was done onto two different plates. One plate containing IPTG was considered as induced and the second plate without IPTG was considered as control. Slow growing colonies upon induction were identified and their growth profile was determined by performing shake flask studies in different media viz. LB, TB and M9.

For the second library under pBAD33, a co-transformation protocol was followed where electrocompetent cells of DH5 $\alpha$  already containing the pNER31 plasmid (with GFP) was prepared by glycerol wash method and electroporated with the ligation mixture containing the genomic library. Primary screening was done on LB agar plates by induction of both the plasmids with IPTG (1 mM final concentration) and arabinose (0.02%, final concentration). Colonies with enhanced fluorescence were picked, and the genomic fragments cloned in the second plasmid (containing the library) which were responsible for enhanced fluorescence, was determined by sequencing. The growth and product profiles were determined for the selected clones by shake flask studies which were done in M9, TB and LB media.

Selected clones which showed enhanced fluorescence upon induction were grown overnight with shaking at 37°C in 3 ml LB tubes. Secondary inoculation was done by adding 100  $\mu$ l of overnight grown culture in 10 ml of LB, TB and M9 medium in 100 ml flasks. After 1.5 hours when the OD<sub>600</sub> of 0.3-0.5 was attained, cultures were induced by adding 1 mM IPTG (final concentration). After induction, the OD<sub>600</sub> was monitored at regular intervals.

For measuring the metabolic activity of non growing cells in minimal media, 200  $\mu$ l of samples were taken at regular intervals post induction. These samples were centrifuged at 13,200 rpm for 2 minutes and the supernatant was stored at 4°C. Glucose levels were estimated by the 3, 5-Dinitrosalicylic acid (DNSA) method.

The expression of GFP was monitored by Varian (Cary Eclipse) Fluorescence Spectrophotometer with an excitation wavelength of 488 nm and emission wavelength of 514 nm.

The specific yield was calculated using the formula Fluorescence/OD = AU/gm DCW (the dry cell weight factor is 0.3 gm/L/1OD600).

### Abbreviations

GFP: Green fluorescent protein; CTAB: Hexadecyltrimethylammonium bromide; DNSA: 3, 5-Dinitrosalicylic acid; CIAP: Calf Intestinal alkaline phosphatase; IPTG: Isopropyl  $\beta$ -D-1-thiogalactopyranoside.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

CG and KJM designed the experiments and wrote the manuscript. CG performed all the experiments. RG performed few experiments. All authors read and approved the final manuscript.

### Acknowledgement

CG gratefully acknowledges CSIR fellowship, Govt. of India. The authors thank Dr. Beckwith (HMS, Boston), Dr. Summers (Cambridge, UK) and Dr. J. Singh (ICGEB, India) for gifting the plasmids pBAD33, pNER31 and pET14b-GFP respectively.

Received: 18 January 2012 Accepted: 11 June 2012

Published: 3 July 2012

### References

1. Bailey JE: Towards a science of metabolic engineering. *Science* 1991, **252**:1668–1675.
2. Stephanopoulos G, Sinskey AJ: Metabolic engineering methodologies and future prospects. *Trends Biotechnol* 1993, **11**:392–396.
3. Stephanopoulos G: Metabolic fluxes and metabolic engineering. *Metab Eng* 1999, **1**:1–11.
4. Stephanopoulos G, Kelleher J: How to make a superior cell. *Science* 2001, **292**:2024–2026.
5. Baudin A, Ozier-Kalogeropoulos O, Denouel A, Lacroute F, Cullin C: A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 1993, **21**:3329–3330.
6. Datsenko KA, Wanner BL: One step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 2000, **97**:6640–6645.
7. Lutz R, Bujard H: Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I-1-2 regulatory elements. *Nucleic Acids Res* 1997, **25**:1203–1210.
8. Mumberg D, Muller R, Funk M: Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 1995, **156**:119–122.
9. Voss I, Steinbuchel A: Application of a KDPG-aldolase gene-dependent addiction system for enhanced production of cyanophycin in *Ralstonia eutropha* strain H16. *Metab Eng* 2006, **8**:66–78.
10. Yong J, Stephanopoulos G: Multi-dimensional gene target search for improving lycopene biosynthesis in *Escherichia coli*. *Metab Eng* 2007, **9**:337–347.
11. Lee SY, Kim HU, Park JH, Park JM, Kim TY: Metabolic engineering of microorganisms: general strategies and drug production. *Drug Discov Today* 2009, **14**:78–88.



12. Chou CP: Engineering cell physiology to enhance recombinant protein production in *Escherichia coli*. *Appl Microbiol Biotechnol* 2007, **76**:521–532.
13. Choi JH, Lee SJ, Lee SY: Enhanced production of insulin-like growth factor I fusion protein in *Escherichia coli* by coexpression of the down-regulated genes identified by transcriptome profiling. *Appl Environ Microbiol* 2003, **69**:4737–4742.
14. March JC, Eiteman MA, Altman E: Expression of an anaplerotic enzyme, pyruvate carboxylase, improves recombinant protein production in *Escherichia coli*. *Appl Environ Microbiol* 2002, **68**:5620–5624.
15. Weikert C, Canonaco F, Sauer U, Bailey JE: Co-overexpression of RspAB improves recombinant protein production in *Escherichia coli*. *Metab Eng* 2000, **2**:293–299.
16. Lütke-Eversloh T, Stephanopoulos G: Combinatorial pathway analysis for improved L-tyrosine production in *Escherichia coli*: identification of enzymatic bottlenecks by systematic gene overexpression. *Metab Eng* 2008, **10**:69–77.
17. Mukherjee KJ, Rowe DC, Watkins NA, Summers DK: Studies of single-chain antibody expression in quiescent *Escherichia coli*. *Appl Environ Microbiol* 2004, **70**:3005–3012.
18. Bailey JA, Sburlati A, Hatzimanikatis V, Lee K, Renner W, Tsai P: Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. *Biotechnol Bioeng* 1996, **52**:109–121.
19. Delgado J, Liao JC: Inverse flux analysis for reduction of acetate excretion in *Escherichia coli*. *Biotechnol Prog* 1997, **13**:361–367.
20. Sonderegger M, Schümperli M, Sauer U: Selection of quiescent *Escherichia coli* with high metabolic activity. *Metab Eng* 2005, **7**:4–9.
21. Srivastava P, Bhattacharaya P, Pandey G, Mukherjee KJ: Overexpression and purification of recombinant human interferon alpha2b in *Escherichia coli*. *Protein Expr Purif* 2005, **41**(2):313–322.
22. Banack T, Clauson N, Ogbaa N, Villar J, Oliver D, Firshein W: Overexpression of the Hda DnaA-related protein in *Escherichia coli* inhibits multiplication, affects membrane permeability, and induces the SOS response. *J Bacteriol* 2005, **187**(24):8507–8510.
23. Chin JW, Cirino PC: Strain engineering strategies for improving whole-cell biocatalysis: engineering *Escherichia coli* to overproduce xylitol as an example. *Meth Mol Biol* 2011, **743**:185–203.
24. Pandhal J, Ow SY, Noirel J, Wright PC: Improving N-glycosylation efficiency in *Escherichia coli* using shotgun proteomics, metabolic network analysis, and selective reaction monitoring. *Biotechnol Bioeng* 2011, **108**(4):902–912.
25. Waegeman H, De Lausnay S, Beauprez J, Maertens J, De Mey M, Soetaert W: Increasing recombinant protein production in *Escherichia coli* K12 through metabolic engineering. *N Biotechnol* 2011, [Epub ahead of print].
26. Bacher A: *Biosynthesis of flavins*, Chemistry and biochemistry of flavoenzymes, Volume 1. Boca Raton, Fla: CRC press; 1991:215–259.
27. Bacher A, Eisenreich K, Kis K, Ladenstein R, Richter G, Scheuring J, Weinkauff S: *Biosynthesis of flavins*. In *Bioorganic chemistry frontiers*, Volume 3. Edited by Dugas H, Schmidtchen FP. Berlin, Germany Springer; 1993:147–192.
28. Raina S, Mabey L, Georgopoulos C: The *Escherichia coli htrP* gene product is essential for bacterial growth at high temperatures: mapping, cloning, sequencing, and transcriptional regulation of htrP. *J Bacteriol* 1991, **173**:5999–6008.
29. Guzman LM, Belin D, Carson MJ, Beckwith J: Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 1995, **177**:4121–4130.

doi:10.1186/1475-2859-11-93

Cite this article as: Ghosh *et al.*: An inverse metabolic engineering approach for the design of an improved host platform for over-expression of recombinant proteins in *Escherichia coli*. *Microbial Cell Factories* 2012 **11**:93.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

